

Data issues and experimental design

First Semester
Senior Seminar

Data

- Record of systematic observations
- How should I structure the observations?
- How many observations are enough?
- What will I do with them?

Types of Measurements

- Nominal/Categorical
 - male or female; control or treatment
- Ordinal/Ranked
 - low, medium, high
- Continuous/Interval
 - 2.9, 4.2, 3.8

Types of Variables

Dependent: The variable(s) you are interested in being able to explain or predict

Independent/Predictor: The variable(s) you might believe affect the value of your dependent variable(s)

Recording Data

- Notebooks (lab and/or field)
- Digital Recorders/Dataloggers
- Data Sheets
- Methods, Metadata

Entering Data

- Enter as soon as possible
- Record metadata
- Use informative file names
- Proof-read & calculate summary statistics
- Exploratory data analysis

Tracking Data

- Log book/log sheet
- Separate files with informative names
- Live links in Excel
- **BACK UP YOUR DATA!**

Analyzing Data

- Summary statistics (mean, median, range, min, max, etc.)
- Variation (standard deviation, standard error, confidence interval)
- Correlation and regression
- T-tests and ANOVAs
- Contingency tables and Chi-square tests
- Logistic regression

Other Concerns

- Independence of Observations
- Confounding factors
- Randomization
- Realism (both nature and range of treatment)
- Controls
- Consistency of application
- Covariates

Where to go for help

- Your mentor and their associates (to start)
- The library
 - Statistics books
 - Methods section of related papers
- Statistical Consulting: consult@stat.columbia.edu